

## **Assessing language skills in adolescents with autism spectrum disorder**

Diane King<sup>1</sup>, Olympia Palikara<sup>2</sup>

1 National Foundation for Educational Research (NFER)

2 School of Education, University of Roehampton

Dr Diane King

National Foundation for Educational Research

The Mere, Upton Park, Slough, Berkshire. SL1 2DQ

E-mail: [d.king@nfer.ac.uk](mailto:d.king@nfer.ac.uk)

### **Corresponding author:**

Dr Olympia Palikara

School of Education, University of Roehampton

Froebel College, Roehampton Lane

London, SW15 5PJ, UK

Email: [Olympia.Palikara@roehampton.a.uk](mailto:Olympia.Palikara@roehampton.a.uk)

**Acknowledgements:** The authors would like to thank those who participated in this study and Professor Julie Dockrell and Professor Morag Stuart for their advice in the development of the experimental narrative tasks.

**Declaration of Conflicting Interests** The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Funding** The author(s) received no financial support for the research, authorship, and/or publication of this article.

## **Assessing language in adolescents with autism spectrum disorder**

### **Abstract**

Language abilities in adolescents with Autism Spectrum Disorders (ASD) are variable and can be challenging to ascertain with confidence. This study aimed to compare and evaluate different forms of language assessment: standardised language testing, narrative analysis and parent/teacher reports. 14 adolescents with ASD and 14 typically developing adolescents matched on age, gender and nonverbal ability were assessed using a number of standardised assessments for receptive and expressive language skills, a standardised narrative test, two experimental narrative assessments and a parent/teacher report measure of pragmatics. The findings were that, although adolescents with ASD scored within the normal range on expressive and receptive language, their performance on narrative tasks revealed difficulties with both structural and evaluative language. It should be noted that both teachers and parents rated the pragmatic language skills of the young people with ASD as significantly lower than those of the typically developing group but parents were more likely than teachers to additionally identify difficulties in speech and syntax. The implications of these results for professionals in terms of assessing the language skills of adolescents with ASD and for the planning of appropriate intervention are discussed.

**Keywords:** autism spectrum disorders, narrative, pragmatic abilities, language assessment, adolescents

## **Assessing language skills in adolescents with autism spectrum disorder**

### **Introduction**

#### *Language skills in autism spectrum disorder*

This study compares different methods of assessing language skills in adolescents with autism spectrum disorder (ASD), a neurodevelopmental disorder characterised by impairments in social interaction, verbal and nonverbal communication and restricted and repetitive behaviours (American Psychiatric Association, 2013). Although the difficulties of individuals with ASD are well documented, there is still much we do not know about the language abilities of this population, in part because of the challenges involved in the assessment of these skills (Hudry et al., 2010). This is an area which merits further research to enable the planning of appropriate programmes of intervention and support.

Language abilities of children and young people with ASD vary greatly, with some individuals' effectively nonverbal (Lord & Paul, 1997) but others displaying no significant impairment on traditional measures of structural language (Williams, Botting & Boucher, 2008). However, it is well established that difficulties with pragmatic language are almost universal in individuals of all ages (Tager-Flusberg, Paul & Lord, 2005). Pragmatic language refers to the appropriate and relevant use of language in context and requires both speaker and listener to have a range of social linguistic skills (Prutting & Kirchner, 1987). Studies report that children and young people with ASD struggle with initiating and turn taking in conversation, topic maintenance, pedantic speech and the use of figurative language. (Diehl, Bennetto & Young, 2005; 2006; Tager-Flusberg, 2000).

*Assessment of language skills in individuals with autism spectrum disorder*

The assessment of language ability in individuals with ASD is problematic. Previous research has shown that on standardised psychometric language measures such as the Clinical Evaluation of Language Fundamentals (CELF; Semel, Wig & Secord, 2006), a test measuring structural aspects of oral language, and the British Picture Vocabulary Scale (BPVS; Dunn & Dunn, 2009), a test assessing receptive vocabulary, high functioning individuals with ASD may score within the normal range, even though it is clear to both professionals and parents that they have significant communication difficulties (Volden & Phillips, 2010; Young et al, 2005). Such tests focus on linguistic structures, syntax and vocabulary in a one to one situation and may not accurately reflect the difficulties individuals experience in using language in an everyday social setting. Not only has this implications for planning targeted programmes of intervention, but there are also consequences for the funding of resources and support (Young et al, 2005). A child or young person who has language scores falling within the average range may be judged not to require additional resources and services (Klin & Volkmar, 2000). In this study, therefore, we also consider alternative but complementary forms of assessment: parental and teacher reports of adolescents' pragmatic language skills and the use of narrative analysis.

*Assessing pragmatic language skills: The Children's Communication Checklist (CCC-2)*

A number of tests have been developed to assess pragmatic language. Of these, the CCC-2 is reported to be the most proficient at identifying pragmatic language impairment (Volden & Phillips, 2010). Carers or professionals are asked to rate the child on a checklist comprising ten scales, four relating to pragmatic language, four assessing structural language and two assessing non-language domains such as autistic features not directly related to language. The

test has not been designed to diagnose ASD, but low scores on pragmatic scales combined with impairment in the non-language domains indicate that the child should be referred for a more detailed clinical assessment. Used on its own, it is thought to have limitations in the identification of autism; Charman et al (2007) questioned its efficacy as a screening instrument for autism, arguing it produces false positives and may overestimate the prevalence of autism.

The use of parent report measures in the CCC-2 overcomes many of the difficulties inherent in formal language testing, but may introduce other biases. It is difficult to ascertain with confidence the validity and reliability of these reports. Furthermore, correlations between parent's ratings of pragmatic language at home and those of professionals indicate some inconsistency across environments. Bishop and Baird (2001) report correlations ranging from 0.30 and 0.58 for individual pragmatic scales and 0.46 for the pragmatic composite. Such discrepancies in children with developmental difficulties are not unusual and have also been reported in assessments of emotional, behavioural and social difficulties (Redmond & Rice, 1998).

#### *Narrative as a tool for assessing language ability in ASD*

There is growing support for the view that the assessment of narrative skills in children and young people with ASD provides useful information about language abilities, beyond that obtained through formal language testing. The construction of a narrative is a complex task, drawing on a number of linguistic, social and cognitive abilities. Individuals must understand and produce language, plot events, use grammatical structures to mark causal and temporal relations and both attribute and understand thoughts, emotions, and intentions (Losh & Capps, 2003). Research suggests that narrative assessment provides an alternative and

ecologically valid method of obtaining information on the language skills of individuals with ASD (Botting, 2002; Manolitsi & Botting, 2011). In a study examining the structural and pragmatic language of children with ASD and Specific Language Impairment (SLI), Manolitsi and Botting found that whilst standardised tests yielded relatively little information which distinguished one group of children from the other, measures on a structured narrative task revealed a number of qualitative differences between the groups. Similar findings are reported by Banney, Harper-Hill and Arnott (2014) who found that, compared to a language matched group of typically developing children, narratives elicited from children with ASD during the Autistic Diagnostic Observation Schedule (ADOS) demonstrated impairments in syntax, use of pronouns and elements of story grammar. However, evidence in this area is not yet conclusive. A study comparing the use of two language tests to identify pragmatic language difficulties in children with ASD, found that the Strong Narrative Assessment Procedure (SNAP; Strong, 1998) failed to differentiate clearly between children with and without ASD (Young et al, 2005).

### *Aims and predictions*

The aim of this study was to compare the performance of adolescents with ASD and their typically developing peers on three different forms of language assessment in order to evaluate the effectiveness of each to identify language difficulties experienced by young people with ASD. Participants were assessed on the CELF IV, the BPVS, the CCC-2 and tasks of narrative analyses across three narrative genres: retelling a story, constructing a fictional story and talking about events. Further, given that previous research has reported low correlations between ratings of parents and professionals of children with developmental difficulties, a secondary aim was to examine differences in parent and teacher ratings on the CCC-2.

## Methods

### *Participants*

The total sample of this study comprised 28 adolescents, 14 with a clinical diagnosis of ASD and 14 typically developing adolescents matched with the ASD group on chronological age and nonverbal ability.

The ASD group consisted of 14 male adolescents with ASD, aged 11–14 years, all with a clinical diagnosis of ASD given prior to the commencement of the study and confirmed by parents and schools. With the exception of one, all had statements of special educational needs identifying an autistic spectrum disorder as their primary need. This is a legal document issued by a local authority in the UK describing the needs of the pupil and how these should be addressed. Three young people attended a mainstream school and the remainder, a specialist ASD school or unit. 14 typically developing adolescents, drawn from two secondary schools were matched with the ASD group on chronological age, gender and nonverbal ability. Mean nonverbal and verbal ability scores for both groups were within the average range for their age. No significant differences were found between the two groups on measures of age, nonverbal or verbal ability (Table 1).

Insert Table 1 about here



## *Materials*

### Standardised tests

All young people were assessed on the following standardised tests:

*Matrices test of the BAS II (British Ability Scales Second Edition, Elliot, Smith, & McCullouch, 1996).*

The matrices test of the BAS II is a measure of nonverbal ability. Participants are presented with a set of patterns where one pattern is incomplete and required to point to the missing piece. The BAS II demonstrates good reliability (.95 for school age children) and validity (.69).

*BPVS II (British Picture Vocabulary Scale, 2nd ed., Dunn & Dunn, 1997).*

The BPVS II is a measure of receptive vocabulary. Participants are shown four line drawings and asked to choose the one that best illustrates a word spoken by the assessor. It has good psychometric properties (reliability .93, validity .76.)

*CELF IV UK (Clinical Evaluation of Language Fundamentals – Fourth Edition, Semel, Wiig, & Secord, 2006).*

The CELF IV consists of core sub tests measuring semantics, syntax and working memory with supplementary subtests for receptive and expressive language and automatic naming ability. Composite scores are calculated from the scores of different combinations of the subtests. The composite scores all have excellent or good reliability scores ranging from .88 to .92.

*Children's Communication Checklist (CCC-2; Bishop, 1998).*

The CCC-2 comprises 70 items divided into nine scales. Two scales assess aspects of language structure (speech and syntax), five assess pragmatic aspects of communication (inappropriate initiation, coherence, stereotyped conversation, use of context, and rapport), and two assess non-linguistic aspects of autistic behaviour (social relationships and interests). Each item contains a statement describing a specific behaviour (e.g., ‘talks repetitively about things that no-one is interested in’) which is rated as ‘definitely applies’, ‘applies somewhat’, ‘does not apply’ or ‘unable to judge’.

*The Expression, Reception and Recall of Narrative Instrument (ERRNI; Bishop, 2004)*

The *ERRNI* is an assessment of spoken narrative skills. Participants are required to look at a series of colour pictures, retell the story, with the pictures in sight and answer a series of comprehension questions. Three of these questions are based on literal information presented in the text with six requiring the participants to make inferences. Reliability scores range from .75-.90.

## Experimental measures

### *Narrative Tasks*

Two narrative tasks developed by King, Dockrell and Stuart (2013; 2014) were used to assess oral narrative language skills.

### *Event narratives*

Participants completed two different event narrative tasks, one designed to elicit a general response and one a specific response. In the general condition, participants were asked six questions encouraging them to recall a general event (e.g. ‘What do people usually do at

Halloween?') In the specific condition, questions were designed to encourage a narrative of a specific personal event (e.g. 'Can you tell me about what you did at Halloween one time?'). Each question was accompanied by a relevant picture. Internal reliability is reported as high (Cronbach's Alpha coefficient of  $> 0.7$ ).

### *Story stems*

In this task, participants were presented with the story stems shown below, with accompanying pictures, and asked to continue the story.

1. The boy ran into the forest. He looked ahead of him and saw a little green man in a spaceship.
2. When the girl climbed up the mountain, she saw, hidden among the trees, a little wooden house covered in snow.

Good internal consistency is reported for this task (Cronbach's Alpha coefficient of  $> 0.7$ ).

### *Procedure*

Each pupil was assessed on three separate occasions individually in a quiet room at school. The order in which the event narratives were presented, the order of the questions, and the order of presentation of the story stems was counterbalanced to control for order effects and possible differences in difficulty. Narratives were recorded and later transcribed and coded.

The event narratives were coded for structural language, evaluative language and enrichment devices using the Systematic Analysis of Language Transcripts (SALT) programme (Miller & Iglesias, 2008). Structural measures included length of narrative (as measured by the number of main body words), mean length of utterance, and the number of different word

roots. Evaluative measures included references to mental states, causal statements and narrative enrichment devices (character speech, negative comments, emphatic markers, and hedges -narrative devices used to distance the speaker, for example, ‘sometimes’, ‘perhaps,’ ‘probably’, ‘might’).

The story stems were coded in the same manner as the event narratives using the SALT programme of analysis but, in addition, a hand coded analysis was undertaken using the narrative scoring scheme (NSS) in the SALT (2008) handbook. Each story was given a score between 0 and 5 on seven categories, 1 being ‘proficient use’ and 5 being ‘minimal use’.

To ensure the reliability of the coding, 10% of the narratives were coded separately by two trained researchers. Inter-rater reliability was found to be high for both the event narrative (.90) and the story stem tasks (.87). The narrative coders were blind to whether the participants were in the ASD group or typically developing group.

Parents and teachers were asked to complete a copy of the CCC-2.

The study was approved by the ethics committee of the UCL Institute of Education and followed British Psychological Society guidelines. Consent for participation was obtained from both parents and schools.

## **Results**

### *Overview of the data and statistical analysis*

Means and standard deviations of the scores of both groups on each of the measures of the administered tests were calculated. Independent *t* tests were conducted to test for significant differences between the groups on the core, expressive and receptive language measures of the CELF IV, the BPVS II, and the narrative assessments. A series of one-way between-

groups ANOVAs were employed to investigate differences between the ASD and comparison group on the scaled scores of each scale of the CCC-2 for both parent and teacher ratings. Differences between teacher and parent ratings of the young people on the CCC-2 scales were examined using a paired samples *t* test. The analysis of the scores on the experimental narrative tasks and the CCC-2 necessitated the use of multiple tests therefore, in order to control for a Type 1 error and, consistent with previous studies examining narrative performance over a number of different measures, we adopted a more stringent alpha level,  $p < 0.01$  (Rumpf et al, 2012). Effect sizes were evaluated with Cohen's *d* using the means and standard deviations of the two groups.

A comparison of the standard scores of both groups on the CELF IV and the BPVS II, as depicted on Table 2, revealed differences between the language profile of the adolescents with ASD and that of the typically developing comparison group. Scores for both groups on the expressive, receptive and core language measures of the CELF IV and receptive vocabulary scores on the BPVS II fell within the average range, but those of the young people with ASD were lower than those of the comparison group. These differences were significant for the core language composite score ( $t(26) = -2.36, p = .03$ , Cohens  $d = 0.89$ ), but not the receptive or expressive language indices of the CELF IV, or the BPVS II.

Insert Table 2 about here

*Comparisons on narrative assessments*

Differences between the two groups on the narrative assessments are shown in Table 3. In the *ERRNI*, standard scores for each group on both measures of the test, ‘ideas’ and ‘mean length of utterance’, fell within the average range but those of the ASD group were significantly lower than on both indices (‘ideas’  $t(26) = -2.44$ ,  $p = .02$ , Cohen’s  $d = 0.92$ ; ‘mean length of utterance’,  $t(26) = -3.89$ ,  $p = .001$  Cohen’s  $d = 1.47$ ).

Insert Table 3 about here

The event narratives of the adolescents with ASD differed from those of the comparison group in both structural and evaluative measures of language (Table 4 and Table 5). They were significantly shorter, (gen ( $t(26) = -2.99$ ,  $p < .01$ , Cohen’s  $d = 1.15$ ; sen ( $t(26) = -2.67$ ,  $p < .01$ , Cohen’s  $d = 1.01$ ), contained proportionally fewer different word roots (gen ( $t(26) = -5.81$ ,  $p < .01$ , Cohen’s  $d = 2.26$ ; sen ( $t(26) = -4.18$ ,  $p < .01$ , Cohen’s  $d = 1.58$ ), had a shorter mean length of utterance (gen ( $t(26) = -5.15$ ,  $p < .01$ , Cohen’s  $d = 2.00$ ; sen ( $t(26) = -5.19$ ,  $p < .01$ , Cohen’s  $d = 1.97$ ), contained fewer causal statements (gen ( $t(26) = -2.91$ ,  $p < .01$ , Cohen’s  $d = 1.15$ ; sen ( $t(26) = -4.60$ ,  $p < .01$ , Cohen’s  $d = 1.8$ ), had fewer references to mental states (gen ( $t(26) = -2.91$ ,  $p < .01$ , Cohen’s  $d = 1.3$ ), and made less use of enrichment devices (gen ( $t(26) = -3.66$ ,  $p < .01$ , Cohen’s  $d = 1.4$ ).

Insert Tables 4 & 5 about here

Tables 6 and 7 show that the mean scores of the ASD group in the story stem task were lower than those of the comparison group on all structural and evaluative measures although differences were significant for just two of the measures: mean length of utterance and number of causal statements (mlu:  $t(26) = -3.12, p < .01$ , Cohen's  $d = 1.18$ ; causal statements:  $t(26) = -2.78, p < .01$ , Cohen's  $d = 1.05$ ). However, the results of the analysis of the NSS codes of the stories showed that stories produced by the ASD group differed significantly from those of the comparison group on a number of aspects of storytelling (character development:  $t(26) = -4.24, p < .01$ , Cohen's  $d = 1.60$ ; references to mental states:  $t(26) = -3.51, p < .01$ , Cohen's  $d = 1.33$ ; referencing:  $t(26) = -4.18, p < .01$ , Cohen's  $d = 1.58$ ; conflict resolution:  $t(26) = -2.98, p < .01$ , Cohen's  $d = 1.13$ ; cohesion:  $t(26) = -3.33, p < .01$ , Cohen's  $d = 1.27$ ; and total narrative score  $t(26) = -3.40, p < .01$ , Cohen's  $d = 1.29$ ).

Insert Tables 6 & 7 about here

#### *Comparisons on the CCC-2 observational checklist*

Table 8 shows that mean scaled scores of parent and teacher CCC-2 checklists for the adolescents with ASD are considerably lower than those of the standardisation sample. Using interpretation guidance from the CCC-2 manual, the scores for this group fall within normal limits on the speech and syntax scales but, on most other scales, are below the 5<sup>th</sup> percentile, suggesting communication problems of clinical significance. Scores below the 6<sup>th</sup> percentile

on the two scales of the SIDC of the CCC-2 indicate a communicative profile characteristic of adolescents with ASD.

Insert Table 8 about here

### *Teacher ratings*

Teacher ratings of the ASD group were significantly lower than those of the comparison group on the *semantics* [ $F(1, 26) = 14.97, p = 0.005$ , Cohen's  $d = 1.46$ ], *coherence* [ $F(1, 26) = 16.76, p = 0.00$ , Cohen's  $d = 1.55$ ], *inappropriate initiation* [ $F(1, 26) = 11.20, p = 0.002$ , Cohen's  $d = 1.26$ ], *stereotyped language* [ $F(1, 26) = 24.48, p = 0.00$ , Cohen's  $d = 1.87$ ], *use of context* [ $F(1, 26) = 44.16, p = 0.00$ , Cohen's  $d = 2.51$ ], *nonverbal communication* [ $F(1, 26) = 59.82, p = 0.00$ , Cohen's  $d = 2.92$ ], *social relations* [ $F(1, 26) = 49.89, p = 0.005$ , Cohen's  $d = 2.67$ ] and *interests'* [ $F(1, 26) = 34.64, p = 0.00$ , Cohen's  $d = 2.23$ ] scales of the CCC-2. No significant differences were found between their ratings of the groups on the *speech* and *syntax* scales.

### *Parent ratings*

Parental ratings of adolescents with ASD on all scales of the CCC-2 were significantly lower than those of the comparison group on all measures (*speech* [ $F(1, 26) = 14.04, p = 0.002$ , Cohen's  $d = 1.42$ ], *syntax* [ $F(1, 26) = 113.73, p = 0.002$ , Cohen's  $d = 1.40$ ], *semantics* [ $F(1, 26) = 25.92, p = 0.00$ , Cohen's  $d = 1.93$ ], *coherence* [ $F(1, 26) = 30.61, p = 0.00$ , Cohen's  $d = 2.09$ ], *inappropriate initiation* [ $F(1, 26) = 47.63, p = 0.00$ , Cohen's  $d = 2.61$ ], *stereotyped language* [ $F(1, 26) = 33.28, p = 0.00$ , Cohen's  $d = 2.18$ ], *use of context* [ $F(1, 26) = 74.11, p =$



0.00, Cohen's  $d = 3.25$ ], *nonverbal communication* [ $F(1, 26) = 62.20, p = 0.00$ , Cohen's  $d = 2.97$ ], *social relations* [ $F(1, 26) = 56.82, p = 0.00$ , Cohen's  $d = 2.84$ ], and *interests* [ $F(1, 26) = 63.15, p = 0.00$ , Cohen's  $d = 2.99$ ].

A secondary aim of the study was to investigate teacher and parent differences in ratings of adolescents on the CCC-2 scales. Results showed that ratings of the parents of the adolescents with ASD on the CCC-2 ( $M = 3.35, SD = 2.24$ ) were significantly lower than teacher's ratings on the *inappropriate initiation* scale ( $M = 6.00, SD = 3.16$ ),  $t(13) = 4.17, p = .001$ , Cohen's  $d = 0.96$  and the *interests* scale ( $M = 2.50, SD = 1.69$ ), ( $M = 5.14, SD = 2.44$ ),  $t(13) = 5.3, p = .000$ , Cohen's  $d = 1.25$ . No significant differences were found between the parent and teacher ratings on any of the other scales. There were no significant differences between parent and teacher ratings of adolescents in the control group on any scale of the CCC-2.

## Discussion

This study compared the performance of adolescents with ASD with that of a matched typically developing group on three forms of language assessment. In line with findings from previous studies (Manolitsi & Botting, 2011; Kjelgaard, & Tager-Flusberg, 2001), we found that the scores of the adolescents with ASD fell within the average range on standardised tests of language ability measuring oral structural language and receptive vocabulary. Scores on the BPVS, the core language composite and the expressive and receptive indexes of the CELF IV, were lower than those of the comparison group but these differences were only significant for the core language composite score. From these findings alone, it may be difficult to argue that the young person requires extra support and funding.

On the other hand, the parent/teacher reports and the analysis of the narrative tasks clearly indicate that adolescents in the ASD group have communication problems of clinical significance. However, there were some interesting differences between the ratings of teachers and parents. Teachers rated adolescents with ASD significantly lower than the comparison group on all scales, apart from speech and syntax, but parent's ratings were significantly lower on all scales, including speech and syntax. This may reflect the fact that teacher's ratings were mostly drawn from observations made in structured classroom situations whereas parent's ratings were more likely to be based on interactions in less structured situations. There is clearly much to be gained from listening to the views of both parents and teachers when assessing young people with ASD (Volden & Phillips, 2010) but it could equally be argued that such information may be biased and unreliable.

The narratives of the adolescents with ASD also differed from those of the comparison group on a number of measures. Three types of narrative tasks were analysed, each yielding different information about the language abilities of the ASD group. Mean scores of both groups on the two measures of the ERRNI fell within the average range but those of the ASD group were significantly lower than those of the comparison group indicating their narratives were shorter and less grammatically complex than those of their typically developing peers. Their event narratives were also significantly shorter and less grammatically complex but additionally contained more limited vocabulary, included fewer reasons and explanations, made fewer references to emotions and thoughts and made less use of linguistic enrichment devices. An analysis of the story stem narratives showed that adolescents with ASD scored lower than the typically developing group on all measures but that differences were significant for just two of these: mean length of utterance and the number of causal statements. However, the NSS analysis rating the story stem narratives from a more 'global' perspective, found that stories of the ASD group differed significantly from those of the

comparison group on aspects related to character development, references to mental states, referencing, conflict resolution and coherence, concurring with findings reported in previous research (Tager-Flusberg & Sullivan, 1995; Losh & Capps, 2003; Deihl et al, 2006).

### ***Educational, research and clinical implications***

The findings of the present study have educational and clinical implications for those working with young people with ASD in education, research and clinical practice. However, they must be interpreted with caution as this is a small study with participants of a limited age range and cognitive ability. Furthermore, whilst tests such as the BPVS II and the CELF IV are judged high on reliability measures and, for the most part, scoring is objective, parents/teacher ratings on the CCC-2 and the coding and analysis of the narratives may be more subjective and open to bias. Moreover, whereas the CCC-2 is relatively easy to administer and score, narrative analysis can be time consuming and complex, although advances in technology are making this form of assessment more manageable. Further, although this study has examined a broad range of language abilities in young people with ASD, it is by no means comprehensive and there are other areas which would merit further exploration. Many studies report impairment or delay in individuals with an ASD in figurative language -the ability to go beyond that which has been explicitly stated (Happe, 1995; Rundblad & Annaz, 2010). Although there is evidence that this may be related to core language skills, syntax and vocabulary (Whyte & Nelson, 2015; Kalandadze, Norbury, Naerland & Naess, 2016), an assessment of figurative language skills may yield useful information for clinicians and educators.

Notwithstanding, we believe that our study raises a number of issues about the assessment of language abilities in young people with an ASD. Our results indicate that scores attained by adolescents with an ASD on commonly used standardised tests of language ability may provide only limited information about the strengths and difficulties they may have with language and communication. The implications of this cannot be underestimated as service providers both in the UK and internationally often regard the scores achieved on these tests as providing them with reliable scientific information for their decisions about the provision of support and resources for individual children (King & King, 2006; Webb & Whitaker, 2012). The assessment of language and communication skills in young people with ASD is complex and the use of additional methods such as an observational checklist and narrative analysis to complement traditional standardised tests may help to provide more detailed information about specific areas of difficulty and aid in the planning of appropriate intervention and support.

**Declaration of Conflicting Interests** The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Funding** The author(s) received no financial support for the research, authorship, and/or publication of this article.

## References

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5<sup>th</sup> ed.). Arlington, VA: American Psychiatric Publishing
- Bishop, D. V. (1998). Development of the Children's Communication Checklist (CCC): A method for assessing qualitative aspects of communicative impairment in children. *Journal of Child Psychology and Psychiatry*, 39(06), 879-891.
- Bishop, D. V. M. (2003). *The Children's Communication Checklist: CCC-2*. London: Harcourt Assessment.
- Bishop, D. V. (2004). *Expression, reception and recall of narrative instrument: ERRNI manual*. Harcourt assessment.
- Bishop, D. V., & Baird, G. (2001). Parent and teacher report of pragmatic aspects of communication: Use of the Children's Communication Checklist in a clinical setting. *Developmental Medicine & Child Neurology*, 43(12), 809-818.
- Botting, N. (2002). Narrative as a tool for the assessment of linguistic and pragmatic impairments. *Child language teaching and therapy*, 18(1), 1-21.
- Charman, T., Baird, G., Simonoff, E., Loucas, T., Chandler, S., Meldrum, D., & Pickles, A. (2007). Efficacy of three screening instruments in the identification of autistic-spectrum disorders. *The British Journal of Psychiatry*, 191(6), 554-559.
- Diehl, J. J., Bennetto, L., & Young, E. C. (2006). Story recall and narrative coherence of high-functioning children with autism spectrum disorders. *Journal of abnormal child psychology*, 34(1), 83-98.
- Dunn, L. M., & Dunn, D. M. (2009). *The British picture vocabulary scale*. GL Assessment Limited.

Dunn, L. M., Dunn, L. M., Whetton, C., & Burley, J. (1997). The British Picture Vocabulary Scale, 2nd ed (Windsor: NFER-Nelson).

Elliott, C. D., Smith, P., & McCulloch, K. (1996). British Ability Scales second edition (BAS II): administration and scoring manual. *London: NFER-Nelson*.

Happé, F. G. (1995). Understanding minds and metaphors: Insights from the study of figurative language in autism. *Metaphor and symbol*, 10(4), 275-295.

Hudry, K., Leadbitter, K., Temple, K., Slonims, V., McConachie, H., Aldred, C., Howlin, P., Charman, T. and the PACT Consortium (2010). Preschoolers with autism show greater impairment in receptive compared with expressive language abilities. *International journal of language & communication disorders*, 45(6), 681-690.

Geurts, H. M., & Embrechts, M. (2008). Language profiles in ASD, SLI, and ADHD. *Journal of Autism and Developmental Disorders*, 38(10), 1931-1943.

Kalandadze, T., Norbury, C., Nærlund, T., & Næss, K. A. B. (2016). Figurative language comprehension in individuals with autism spectrum disorder: A meta-analytic review. *Autism*, 1362361316668652.

King, D., Dockrell, J. E., & Stuart, M. (2013). Event narratives in 11–14 year olds with autistic spectrum disorder. *International Journal of Language & Communication Disorders*, 48(5), 522-533.

King, D., Dockrell, J., & Stuart, M. (2014). Constructing fictional stories: A study of story narratives by children with autistic spectrum disorder. *Research in developmental disabilities*, 35(10), 2438-2449.

King, M. and King, D. (2006). How the Law defines the special educational needs of autistic children. *Child and Family Law Quarterly*, 18, No 1, 23-42.

Kjelgaard, M. M., & Tager-Flusberg, H. (2001). An investigation of language impairment in autism: Implications for genetic subgroups. *Language and cognitive processes*, 16(2-3), 287-308.

Lord, C., & Paul, R. (1997). Language and communication in autism. *Handbook of autism and pervasive developmental disorders*, 2, 195-225.

Losh, M., & Capps, L. (2003). Narrative ability in high-functioning children with autism or Asperger's syndrome. *Journal of autism and developmental disorders*, 33(3), 239-251.

Manolitsi, M., & Botting, N. (2011). Language abilities in children with autism and language impairment: using narrative as a additional source of clinical information. *Child Language Teaching and Therapy*, 27(1), 39-55.

Phelps-Terasaki, D., & Phelps-Gunn, T. (1992). Test of Pragmatic Language. Austin, TX: Pro-Ed.

Rundblad, G., & Annaz, D. (2010). The atypical development of metaphor and metonymy comprehension in children with autism. *Autism*, 14(1), 29-46.

Stein, N. L., & Albro, E. R. (1997). Of Goal-Structured Knowledge in Telling Stories. *Narrative development: Six approaches*, 5.

Tager-Flusberg, H. (2000). Language and understanding minds: Connections in autism. *Understanding other minds: Perspectives from developmental cognitive neuroscience*, 2, 124-149.

Tager-Flusberg, H., Paul, R., & Lord, C. (2005). Language and communication in autism. *Handbook of autism and pervasive developmental disorders*, 1, 335-364.

Volden, J., & Phillips, L. (2010). Measuring pragmatic language in speakers with autism spectrum disorders: Comparing the Children's Communication Checklist—2 and the Test of Pragmatic Language. *American Journal of Speech-Language Pathology*, 19(3), 204-212

Young, E. C., Diehl, J. J., Morris, D., Hyman, S. L., & Bennetto, L. (2005). The use of two language tests to identify pragmatic language problems in children with autism spectrum disorders. *Language, Speech, and Hearing Services in Schools, 36*(1), 62-72.

Webb, J. & Whitaker, S. (2012). Defining learning disability. *The Psychologist, 25*, 440-443.

Williams, D., Botting, N., & Boucher, J. (2008). Language in autism and specific language impairment: Where are the links? *Psychological bulletin, 134*(6), 944.

Whyte, E. M., & Nelson, K. E. (2015). Trajectories of pragmatic and nonliteral language development in children with autism spectrum disorders. *Journal of communication disorders, 54*, 2-14.



Table 1. *Characteristics of ASD and Chronological Age Match Group (CM)*

Group	ASD <i>n=14</i> <i>M</i> ( <i>SD</i> )	CM <i>n=14</i> <i>M</i> ( <i>SD</i> )	Significance
Age in months	149.14 (10.37)	150.78 (9.69)	ASD = CM $t(26) = -.43$ $p = .67$
Matrices (BAS II) Nonverbal ability standard score	103.43 (14.59)	101.28 (11.97)	ASD = CM $t(26) = .42$ $p = .67$
BPVS II standard score	101.64 (20.76)	111.00 (17.17)	ASD = CM $t(26) = 1.30$ $p = .20$

Table 2. Means and standard deviations for CELF IV standard scores on language indexes and BPVS II standard scores for ASD and chronological age match groups (CM).

Test	ASD <i>N</i> = 14	CM <i>N</i> = 14	Significance
	<i>M</i> ( <i>SD</i> )	<i>M</i> ( <i>SD</i> )	
Core Language	83.78 (24.16)	101.42 (14.06)	ASD < CM <i>t</i> (26) = -2.36 <i>p</i> = .03
Receptive Language	89.10 (21.30)	98.70 (8.60)	ASD = CM <i>t</i> (26) = -2.01 <i>p</i> = .06
Expressive Language	90.30 (20.20)	101.30 (11.20)	ASD = CM <i>t</i> (26) = -1.91 <i>p</i> = .07
BPVS II standard score	101.64 (20.76)	111.00 (17.17)	ASD = CM <i>t</i> (26) = 1.30 <i>p</i> = .20

Table 3. Means and standard deviations for ERRNI standard scores for ASD and chronological age match groups (CM).

Test	ASD <i>N</i> = 14	CM <i>N</i> = 14	Significance
	<i>M</i> ( <i>SD</i> )	<i>M</i> ( <i>SD</i> )	
Ideas initial storytelling	91.93 (19.55)	108.07 (15.12)	ASD < CM $t(26) = -2.44$ $p = .02$
Mean length of utterance	88.71 (16.81)	112.14 (14.96)	ASD < CM $t(26) = -3.89$ $p = .001$

Table 4: *Means and standard deviations for structural measures in General and Specific Event Narratives for ASD and age match groups*

Measure	ASD <i>N</i> = 14		Age match <i>N</i> = 14		Significance
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
GEN					
No. Of Utterances	28.84	9.54	23.42	6.92	<i>t</i> (26)=1.69 p > .05
No. of Main Body Words	153.15	103.78	281.07	116.98	<i>t</i> (26)=2.99 p < .01
MLU in Morphemes	5.52	2.47	12.29	4.09	<i>t</i> (26)=5.15 p < .01
No. of Word Roots	2.32	0.99	5.00	1.35	<i>t</i> (26)=5.81 p < .01
(proportion of utterances)					
SEN					
No. Of Utterances	35.28	10.85	37.28	15.04	<i>t</i> (26)=-.40 p > .05
No. of Main Body Words	241.28	148.50	433.92	224.35	<i>t</i> (26)=2.67 p < .01
MLU in Morphemes	6.50	2.93	11.23	1.71	<i>t</i> (26)=5.19 p < .01
No. of Word Roots	2.73	1.14	4.22	0.68	<i>t</i> (26)=4.18 p < .01

Table 5: *Mean scores and standard deviations of proportional scores for evaluative measures in General and Specific event narratives for all groups*

Measure	Condition	Statistic	ASD	CM	Significance
Mental States	GEN	<i>M</i>	0.18	0.45	$t(26)=-3.32$
		<i>SD</i>	0.18	0.23	$p < .01$
	SEN	<i>M</i>	0.21	0.33	$t(26)=-1.80$
		<i>SD</i>	0.16	0.17	$p > .05$
Causal Statements	GEN	<i>M</i>	0.06	0.22	$t(26)=-2.91$
		<i>SD</i>	0.08	0.18	$p < .01$
	SEN	<i>M</i>	0.05	0.18	$t(26)=-4.60$
		<i>SD</i>	0.06	0.08	$p < .01$
Total Evaluative Devices	GEN	<i>M</i>	0.64	1.63	$t(26)=-3.66$
		<i>SD</i>	0.62	0.75	$p < .01$
	SEN	<i>M</i>	0.74	1.22	$t(26)=-2.31$
		<i>SD</i>	0.54	0.53	$p < .05$

GEN – General Event Narrative

SEN – Specific Event Narrative

Table 6. *Results for analysis of Story Stem Structural and Evaluative Measures for Autistic Spectrum Disorder (ASD), Language Match (LM) and Chronological Age Match (CM) Groups.*

Measure	ASD <i>M</i> ( <i>SD</i> )	CM <i>M</i> ( <i>SD</i> )	Significance
No. of main body words	150.78 (149.93)	236.00 (170.76)	ASD=CM $t(26)=-1.40$ , $p > .05$
MLU	13.33 (3.16)	17.81 (4.32)	ASD<CM $t(26)=-3.12$ , $p < .01$
No. of word roots	10.07 (2.11)	12.20 (3.37)	ASD=CM $t(26)=-2.00$ , $p < .06$
References to mental states	.35 (.31)	.61 (.34)	ASD= CM $t(26)=-2.07$ , $p < .05$
Causal statements	.22 (.22)	.54 (.37)	ASD<CM $t(26)=-2.78$ , $p < .01$
Total evaluative devices	1.07 (.96)	1.91 (.87)	ASD=LM=CM $t(26)=-2.42$ , $p < .02$

Table 7. *Results of analysis of NSS Codes for Story Stems for Autistic Spectrum Disorders (ASD), Language Match (LM) and Chronological Age Match (CAM) Groups.*

Measure	ASD <i>M</i> ( <i>SD</i> )	CM <i>M</i> ( <i>SD</i> )	Significance
Introduction	4.28 (2.33)	6.21 (1.47)	ASD<CM $t(26)=-2.61$ , $p < .01$
Character Development	3.71 (1.81)	6.21 (1.25)	ASD<CM $t(26)=-4.24$ , $p < .01$
Mental States	2.64 (2.53)	5.64 (1.94)	ASD<CM $t(26)=-3.51$ , $p < .01$
Referencing	4.14 (1.74)	6.35 (.92)	ASD<CM $t(26)=-4.18$ , $p < .01$
Conflict Resolution	2.78 (2.42)	5.00 (1.35)	ASD<CM $t(26)=-2.98$ , $p < .01$
Coherence	3.78 (1.96)	6.00 (1.51)	ASD<CM $t(26)=-3.33$ , $p < .01$
Conclusion	2.50 (2.17)	3.85 (1.51)	ASD=CM $t(26)=-1.91$ , $p > .05$
Total narrative score	24.00 (12.60)	38.07 (8.94)	ASD<CM $t(26)=-3.40$ , $p < .01$

Table 8. *Mean (sd) scaled scores CCC-2: Parental and teacher ratings of children with ASD, age matched control group and standardisation sample*

	<i>ASD group</i> <i>Parent</i> <i>rating</i> <i>(n = 14)</i>	<i>ASD group</i> <i>Teacher</i> <i>rating</i> <i>(n = 14)</i>	<i>Control</i> <i>group</i> <i>Parent</i> <i>rating</i> <i>(n = 14)</i>	<i>Control</i> <i>group</i> <i>Teacher</i> <i>rating</i> <i>(n = 14)</i>	<i>Standardisation</i> <i>sample</i>
Speech	6.50 (3.91)	7.50 (3.67)	10.57 (1.09)	8.78 (3.17)	10.75 (2.40)
Syntax	5.57 (4.23)	7.07 (4.23)	10.21 (2.00)	9.14 (3.08)	11.20 (1.24)
Semantics	3.64 (3.71)	4.35 (3.39)	10.07 (2.92)	9.57 (3.73)	11.50 (2.84)
Coherence	4.00 (2.77)	4.85 (3.42)	10.00 (2.96)	9.85 (3.03)	11.65 (2.35)
In initiation	3.35 (2.24)	6.00 (3.16)	10.85 (3.39)	10.00 (3.16)	9.85 (3.03)
Ster language	3.42 (2.44)	4.64 (2.92)	9.43 (3.03)	10.14 (2.95)	10.90 (2.63)
Use context	1.64 (1.94)	2.42 (2.53)	10.71 (3.43)	10.21 (3.58)	10.85 (2.39)
Non comm.	2.07 (3.11)	2.28 (2.23)	10.43 (2.47)	9.50 (2.68)	11.70 (2.18)
Soc relations	1.14 (1.66)	0.78 (1.12)	9.00 (3.53)	8.28 (3.81)	11.20 (2.04)
Interests	2.50 (1.71)	5.14 (2.44)	10.78 (3.51)	11.57 (3.27)	10.50 (3.32)
GCC Score	30.21 (19.50)	39.14 (20.29)	82.28 (18.31)	77.21(22.10)	88.40
SIDC score	10.64 (10.36)	9.57 (10.90)	0.21 (7.37)	2.00 (8.06)	

GCC – General Communication Composite

SIDC - Social Interaction Deviance Composite